

# TWO EXAMPLES OF NON STRICTLY CONVEX LARGE DEVIATIONS

STEFANO DE MARCO, ANTOINE JACQUIER, AND PATRICK ROOME

**ABSTRACT.** We present two examples of a large deviations principle where the rate function is not strictly convex. This is motivated by a model used in mathematical finance (the Heston model), and adds a new item to the zoology of non strictly convex large deviations. For one of these examples, we show that the rate function of the Cramér-type of large deviations coincides with that of the Freidlin-Wentzell when contraction principles are applied.

## 1. INTRODUCTION

The Gärtner-Ellis theorem is a key result in the theory of (finite-dimensional) large deviations. Extending the results of Cramér [10] for sequences of random variables not necessarily independent and identically distributed (iid), it provides a large deviations framework solely based on the knowledge of the cumulant generating function (cgf) of the sequence. The key assumptions are that the pointwise (rescaled) limit of these cgf satisfies some convexity property and becomes steep at the boundaries of its effective domain; this in turns implies that the rate function governing the large deviations, defined as the convex dual, is also convex. Note that, by definition of the convex dual, essential smoothness of the limiting cgf implies strict convexity of the rate function. When these assumptions are not met, large deviations (potentially with non strictly convex rate function) may or may not hold; the classical example [12, Remark (d), page 45] is that of the sequence  $(Z_n)_{n \in \mathbb{N}}$  distributed as exponential random variables with parameter  $n$ . It is immediate to see that  $\Lambda(u) := \lim_{n \uparrow \infty} n^{-1} \log \mathbb{E}(e^{nuZ_n}) = 0$  if  $u < 1$  and is infinite otherwise. Clearly, the function  $\Lambda$  is not essentially smooth (see Definition A.2), and the assumptions of the Gärtner-Ellis theorem, recalled in Appendix A, are violated; nevertheless, a simple computation reveals that the conclusion of the latter still holds, namely that a large deviations principle exists, with speed  $n$  and rate function  $\Lambda^*(x) := \sup_u (ux - \Lambda(u)) = x$  if  $x \geq 0$ , and infinity otherwise; clearly here, strict convexity of  $\Lambda^*$  does not hold. Dembo and Zeitouni [11] and Bryc and Dembo [7]—in the context of quadratic functionals of Gaussian processes—have proposed a way to bypass this absence of essential smoothness (of  $\Lambda$ ) / strict convexity (of  $\Lambda^*$ ) issue by making the change of measure (key tool in the proof of the Gärtner-Ellis theorem) time-dependent. More recently, O’Brien [27] and Comman [8] have strengthened this theorem, by partially relaxing the steepness and convexity assumptions. In a general infinite-dimensional setting, Bryc’s Theorem [6] (see also [12, Chapter 4.4]), or ‘Inverse Varadhan’s lemma’, allows for large deviations with non convex rate functions. One of the hypotheses this theorem relies on is an exponential tightness requirement on the family of random variables under consideration, which is not always easy to verify. However, several examples have been dug out which do not fall into this framework, such as in the setting of random walks with

---

*Date:* April 19, 2016.

*2010 Mathematics Subject Classification.* 60F10.

*Key words and phrases.* large deviations, non-convex rate function, Gärtner-Ellis, stochastic processes.

AJ acknowledges financial support from the EPSRC First Grant EP/M008436/1.

interface [16], occupation measures of Markov chains [23], the on/off Weibull sojourn process [15], or m-variate von Mises statistics [17].

Motivated by recent developments on large deviations in mathematical finance (see in particular [9, 13, 20], and the excellent review paper [29]), we study the small-time behaviour of the solution of the Feller stochastic differential equation (and an integral version of it) when the starting point is null. The absence of Lipschitz continuity of the diffusion coefficient and the degenerate starting condition make it not amenable to the classical Freidlin-Wentzell framework, and the absence of strict convexity of the limiting moment generating function violates the Gärtner-Ellis assumptions. It turns out that a large deviations principle however holds, and one can furthermore reconcile the pathwise large deviations to the marginal (Gärtner-Ellis one) by contraction. We believe this provides a nice example of non-strictly-convex large deviations principle in the context of continuous-time stochastic processes. It also sheds light on the importance of the starting point of the SDE being null, as opposed to the non-zero case where the Gärtner-Ellis theorem applies directly (see [18]). In Section 2, we present the model and state the large deviations results as time tends to zero; we also establish the connection with the Freidlin-Wentzell analysis via the contraction principle. The proofs of the main results are gathered in Section 3.

**Notations:** For a set  $G \subset \mathbb{R}^n$ , we shall denote by  $G^o$  and  $\overline{G}$  its respective interior and closure in  $\mathbb{R}^n$ .

## 2. MAIN RESULTS

We consider here the following system of stochastic differential equations:

$$(2.1) \quad \begin{aligned} dX_t &= -\frac{1}{2}V_t dt + \sqrt{V_t} dW_t, & X_0 &= 0, \\ dV_t &= (a + bV_t)dt + \xi\sqrt{V_t} dZ_t, & V_0 &= 0, \\ d\langle W, Z \rangle_t &= \rho dt, \end{aligned}$$

where  $a, \xi > 0$ ,  $b < 0$ ,  $|\rho| < 1$  and  $(W_t)_{t \geq 0}$  and  $(Z_t)_{t \geq 0}$  are two standard Brownian motions. We stress the importance of the parameter  $a$  to be strictly positive; otherwise, the process  $V$ , starting from zero, would just remain null, and the unique solution of (2.1) would simply be the two-dimensional zero process. We shall often make use of the notations  $\bar{\rho} := \sqrt{1 - \rho^2}$  and  $\mu := 2a/\xi^2$ . The SDE for the variance process  $V$  has a unique strong solution by the Yamada-Watanabe conditions [24, Proposition 2.13, page 291]. We further assume that  $\mu > 1$ , which ensures that the origin is unattainable for strictly positive times. Define now the following functions:

$$(2.2) \quad \Lambda_X^*(x) = (u_- \mathbf{1}_{\{x < 0\}} + u_+ \mathbf{1}_{\{x \geq 0\}}) x,$$

$$(2.3) \quad \Lambda_V^*(x) = \begin{cases} 2x/\xi^2, & \text{if } x \geq 0, \\ +\infty, & \text{if } x < 0. \end{cases}$$

for all  $x \in \mathbb{R}$ , where the two real numbers  $u_-$  and  $u_+$  read

$$(2.4) \quad \begin{aligned} u_- &:= \frac{2}{\xi\bar{\rho}} \arctan\left(\frac{\bar{\rho}}{\rho}\right) \mathbf{1}_{\{\rho < 0\}} - \frac{\pi}{\xi} \mathbf{1}_{\{\rho = 0\}} + \frac{2}{\xi\bar{\rho}} \left( \arctan\left(\frac{\bar{\rho}}{\rho}\right) - \pi \right) \mathbf{1}_{\{\rho > 0\}}, \\ u_+ &:= \frac{2}{\xi\bar{\rho}} \arctan\left(\frac{\bar{\rho}}{\rho}\right) \mathbf{1}_{\{\rho > 0\}} + \frac{\pi}{\xi} \mathbf{1}_{\{\rho = 0\}} + \frac{2}{\xi\bar{\rho}} \left( \arctan\left(\frac{\bar{\rho}}{\rho}\right) + \pi \right) \mathbf{1}_{\{\rho < 0\}}. \end{aligned}$$

Note that  $u_-$  (resp.  $u_+$ ) is a decreasing (resp. decreasing) function of  $\rho$  and maps the interval  $(-1, 1)$  to  $(-\infty, -2/\xi)$  (resp.  $(2/\xi, +\infty)$ ). We shall use the subscript/superscript M to represent the quantities related to  $X$  or to  $V$ . For instance  $\Lambda_M^*$  represent  $\Lambda_X^*$  or  $\Lambda_V^*$ . We also denote  $\mathcal{K}_X := \mathbb{R} \setminus \{0\}$  and  $\mathcal{K}_V := (0, \infty)$ .

**2.1. Large deviations results.** The main result of this paper is the following theorem, which provides an example of a sequence of random variables for which the limiting logarithmic cumulant generating function is zero (on its effective domain) but a large deviations principle still holds. This is to be compared to the Gärtner-Ellis theorem [12, Theorem 2.3.6] which requires this limiting function to be steep at the boundaries of its effective domain. As highlighted in the proof, understanding the pointwise limit of the (rescaled) cumulant generating function does not suffice any longer, and its higher-order behaviour is needed to prove large deviations.

**Theorem 2.1.** *Whenever  $M = X$  or  $M = V$ , the family  $(M_t)_{t \geq 0}$  satisfies a LDP with speed  $t$  and rate function  $\Lambda_M^*$  as  $t$  tends to zero.*

A more in-depth analysis reveals a more precise behaviour of the small-time probabilities, which take the following form as  $t$  tends to zero:

$$(2.5) \quad \mathbb{P}(M_t \geq x) = \begin{cases} 1 - C(x)t^{1-\mu} \exp\left(-\frac{\Lambda_M^*(x)}{t}\right) (1 + \mathcal{O}(t)), & \text{if } x < 0, \\ C(x)t^{1-\mu} \exp\left(-\frac{\Lambda_M^*(x)}{t}\right) (1 + \mathcal{O}(t)), & \text{if } x > 0, \end{cases}$$

for all  $x \in \mathcal{K}_M$ , for some (smooth) strictly positive function  $C$ . This analysis is based on the so-called theory of sharp large deviations, developed in [4, 5], and used in [3, 20, 21] for diffusion processes and statistical estimators thereof. It is based on refinements of the Gärtner-Ellis theorem in the case where the limiting cgf is not steep at the boundary; these refinements, using a time-dependent change of measure, were introduced in [7] and [11]. For in general  $1 - \mu \neq \frac{1}{2}$ , we incidentally note that the factor  $t^{1-\mu}$  here is not in line with the classical factor  $t^{1/2}$  found in heat kernel expansions for (the tail of the cumulative distribution function of) elliptic diffusions.

For the process  $V$ , such an analysis, based on the moment generating function, is not really required, since the density of  $V_t$  is known explicitly for each  $t \geq 0$ . Indeed, according to (3.2) below, its moment generating function reads

$$\mathbb{E}(e^{uV_t}) = \left(1 + \frac{u\xi^2}{2b} (1 - e^{bt})\right)^{-\mu} = \left(\frac{\lambda_t}{\lambda_t - u}\right)^\mu,$$

where  $\lambda_t := -\frac{2b}{\xi^2(1-e^{bt})}$ , for all  $u < \lambda_t$ . Hence  $V_t$  is distributed as a Gamma random variable with shape  $\mu$  and rate  $\lambda_t$ . Therefore, for any  $x > 0$ ,

$$\begin{aligned} \mathbb{P}(V_t \geq x) &= \frac{\lambda_t^\mu}{\Gamma(\mu)} \int_x^{+\infty} z^{\mu-1} e^{-\lambda_t z} dz = \frac{1}{\Gamma(\mu)} \int_{\lambda_t x}^{+\infty} z^{\mu-1} e^{-z} dz = \frac{(\lambda_t x)^{\mu-1} e^{-\lambda_t x}}{\Gamma(\mu)} \left[1 + \mathcal{O}\left(\frac{1}{\lambda_t x}\right)\right] \\ &= \frac{\exp\left(\frac{bx}{\xi^2}\right)}{\Gamma(\mu)} \left(\frac{2x}{\xi^2}\right)^{\mu-1} t^{1-\mu} \exp\left(-\frac{2x}{\xi^2 t}\right) [1 + \mathcal{O}(t)], \end{aligned}$$

by the properties of the complementary incomplete Gamma function [28, Chapter 4, Section 2.1], and since  $\lambda_t = \frac{2}{\xi^2 t} - \frac{b}{\xi^2} + \mathcal{O}(t)$ ; therefore (2.5) follows for  $M = V$ . In the proof of Theorem 2.1 below, we will however keep the notation  $M$ , standing for both  $X$  and  $V$ , in order to highlight the fact that an analogous proof holds in both cases.

**2.2. Intuitions from Freidlin-Wentzell analysis.** The proof of Theorem 2.1 will be given in Section 3. In this section, we wish to illustrate how piece-wise linear rate functions such as (2.3) can arise from sample-path large deviations. In order to simplify the framework, consider here the solution  $V$  of the equation obtained

from (2.1) by setting  $V_0 = v_0 > 0$  and  $a = b = 0$ . Setting  $V_t^\varepsilon := V_{\varepsilon^2 t}$  for  $\varepsilon > 0$ , the process  $(V_t^\varepsilon)_{t \geq 0}$  is the (weak) solution of the equation

$$(2.6) \quad dV_t^\varepsilon = \varepsilon \xi \sqrt{V_t^\varepsilon} dZ_t, \quad V_0^\varepsilon = v_0 > 0.$$

Pathwise large deviations (as  $\varepsilon$  tends to zero) for the solution of this SDE fall outside the scope of the classical Freidlin-Wentzell framework (as presented in [12, Chapter 5.6]) since the diffusion coefficient lacks the required global Lipschitz continuity property. Based on properties of Bessel processes, Donati-Martin et al. in [14] proved that, for every  $T > 0$ , the process  $(V_t^\varepsilon)_{t \in [0, T]}$  does satisfy a large deviations principle on the path space  $\mathcal{C}_T = C([0, T]; \mathbb{R}_+)$  of non-negative continuous functions, with speed  $\varepsilon^2$  and rate function  $I_T$  given by

$$(2.7) \quad I_T(\varphi) = \begin{cases} \frac{1}{2} \int_0^T \frac{\dot{\varphi}_t^2}{\xi^2 \varphi_t} \mathbf{1}_{\{\varphi_t > 0\}} dt, & \text{if } \varphi \in \mathcal{C}_T \text{ is absolutely continuous and } \varphi_0 = v_0 \\ +\infty, & \text{otherwise,} \end{cases}$$

where one sets  $y^{-1} \mathbf{1}_{\{y > 0\}} = 0$  when  $y = 0$ . More precisely, this means that the estimates

$$(2.8) \quad - \inf_{\varphi \in G^o} I_T(\varphi) \leq \liminf_{\varepsilon \downarrow 0} \varepsilon^2 \log \mathbb{P}(V^\varepsilon \in G) \leq \limsup_{\varepsilon \downarrow 0} \varepsilon^2 \log \mathbb{P}(V^\varepsilon \in G) \leq - \inf_{\varphi \in \overline{G}} I_T(\varphi)$$

hold for every Borel set  $G \subset \mathcal{C}_T$ . By the contraction principle [12, Theorem 4.2.1], the path estimates (2.8) induce a LDP on  $\mathbb{R}$  for the random variable  $V_1^\varepsilon \triangleq V_{\varepsilon^2}$ , where the rate function is now given by

$$(2.9) \quad \Lambda^*(x) := \inf \{I_1(\varphi) : \varphi \in \mathcal{C}_1, \varphi_0 = v_0, \varphi_1 = x\}.$$

This means that the sequence  $(V_t)_{t \geq 0}$  satisfies a LDP with speed  $t$  as  $t$  tends to zero, namely for every Borel set  $A \subset \mathbb{R}$ ,

$$- \inf_{x \in A^o} \Lambda^*(x) \leq \liminf_{t \downarrow 0} t \log \mathbb{P}(V_t \in A) \leq \limsup_{t \downarrow 0} t \log \mathbb{P}(V_t \in A) \leq - \inf_{x \in \overline{A}} \Lambda^*(x).$$

**Proposition 2.2.** *The rate function  $\Lambda^*$  in (2.9) reads*

$$\Lambda^*(x) = \begin{cases} \frac{2}{\xi^2} (\sqrt{x} - \sqrt{v_0})^2, & \text{if } x \geq 0, \\ +\infty, & \text{if } x < 0. \end{cases}$$

For every  $v_0 > 0$ ,  $\Lambda^*$  is a strictly convex function on the positive real axis ( $\partial_{xx} \Lambda^*(x) = \sqrt{v_0}/(2x^{3/2})$ ), converging pointwise to the affine function  $\Lambda_V^*(x) = 2x/\xi^2$  given in (2.3) as  $v_0$  tends to zero. Note we are not claiming here (with the choice  $a = 0$  we made at the beginning of this section) that  $\Lambda_V^*$  is the rate function for  $V_1^\varepsilon$  in (2.6) when  $v_0 = 0$ : of course, in this case the unique solution to (2.6) is the identically null process  $V^\varepsilon \equiv 0$  (hence, the family  $V_1^\varepsilon$  satisfies a LDP with the trivial rate function  $\bar{I}(\varphi) = 0$  if  $\varphi \equiv 0$ , and  $\bar{I}(\varphi) = +\infty$  otherwise).

*Proof of Proposition 2.2.* Let  $AC_+([0, 1])$  denote the set of absolutely continuous functions on  $[0, 1]$ . If  $x < 0$ , then by definition of  $I_1$ , one has  $I_1(\varphi) = +\infty$  for any  $\varphi$  such that  $\varphi_1 = x$ . Then assume  $x \geq 0$ , and consider  $\varphi \in \mathcal{C}_1$  such that  $\varphi_1 = x$  and  $I(\varphi) < +\infty$ . By the superposition principle (or the chain rule for absolutely continuous functions, see [26, Theorem 3.68]), the function  $\psi \equiv \sqrt{\varphi}$  is absolutely continuous on every interval contained in the open set  $\{\varphi > 0\}$ , with derivative almost surely equal to  $\frac{\dot{\varphi}_t}{2\sqrt{\varphi_t}} \in L^2([0, 1])$ . On  $\{\varphi = 0\}$  one has  $\psi \equiv 0$ ,

therefore  $\dot{\psi}_t = 0$  for every accumulation point of  $\{\varphi = 0\}$  (the isolated points form a finite subset of  $[0, 1]$ ). In summary, it follows from [26, Corollary 3.26] that  $\psi \in AC_+([0, 1])$ , and that

$$\int_0^1 \frac{\dot{\varphi}_t^2}{\varphi_t} \mathbf{1}_{\{\varphi_t > 0\}} dt = 4 \int_{\{\varphi > 0\}} \dot{\psi}_t^2 dt = 4 \int_0^1 \dot{\psi}_t^2 dt.$$

Conversely, let  $\psi \in AC_+([0, 1])$  be such that  $\dot{\psi} \in L^2([0, 1])$ , and set  $\varphi \equiv \psi^2$ ; as the composition of a  $C^1$  function and an absolutely continuous one,  $\varphi$  also belongs to  $AC_+([0, 1])$  and  $\dot{\varphi}_t = 2\psi_t \dot{\psi}_t = 2\sqrt{\varphi_t} \dot{\psi}_t$  a.s. Therefore,

$$\begin{aligned} \Lambda^*(x) &= \inf \left\{ \frac{1}{2\xi^2} \int_0^1 \frac{\dot{\varphi}_t^2}{\varphi_t} \mathbf{1}_{\{\varphi_t > 0\}} dt : \varphi \in AC_+([0, 1]) \text{ and } \varphi_0 = v_0, \varphi_1 = x \right\} \\ (2.10) \quad &= \inf \left\{ \frac{2}{\xi^2} \int_0^1 \dot{\psi}_t^2 dt : \psi \in AC_+([0, 1]) \text{ and } \psi_0 = \sqrt{v_0}, \psi_1 = \sqrt{x} \right\}. \end{aligned}$$

It is well known that the last problem is solved by the straight line  $\psi_t^* \equiv \sqrt{v_0} + t(\sqrt{x} - \sqrt{v_0})$ . Substitution into (2.10) yields  $\Lambda^*(x) = \frac{2}{\xi^2} \int_0^1 (\dot{\psi}_t^*)^2 dt = \frac{2}{\xi^2} (\sqrt{x} - \sqrt{v_0})^2$ , and the proposition follows.  $\square$

Returning to the small-time problem for the solution  $V_t$  of (2.1), now set  $V_t^\varepsilon := V_{\varepsilon^2 t}$ , which satisfies

$$(2.11) \quad dV_t^\varepsilon = \varepsilon^2(a + bV_t^\varepsilon)dt + \varepsilon\xi\sqrt{V_t^\varepsilon}dZ_t, \quad V_0^\varepsilon = 0.$$

To our knowledge, large deviations for the solution to (2.11) are not covered by the existing literature (in [14], the situation where the drift  $a + bV$  is independent of  $\varepsilon$  is considered). We leave it to future research to prove that a pathwise LDP holds for the solution to (2.11) with a rate function analogous to (2.7). Let us point out here that a pathwise LDP for the solution of the SDE

$$d\tilde{V}_t^\varepsilon = \varepsilon^2 a(V_t^\varepsilon)dt + \varepsilon\xi\sqrt{\tilde{V}_t^\varepsilon}dZ_t$$

starting at  $\tilde{V}_0^\varepsilon = \varepsilon^2 v_0$ , was proven in [9, Theorem 1.1]. Comparing with (2.11), note that the initial condition is strictly positive, but tends to zero as  $\varepsilon$  tends to zero; the condition  $b = 0$  is assumed, but the coefficient  $a$  is allowed to be a bounded and Lipschitz continuous function with values in  $\mathbb{R}_+$ .

**Remark 2.3.** The resulting rate function for the process  $\tilde{V}^\varepsilon$  is identical to  $I_T(\varphi)$  given in (2.7), albeit now with initial condition  $\varphi_0 = 0$ , for any choice of  $a$  within this class. Following analogous arguments to the proof of Proposition 2.2, the contraction principle applied to  $I_T$  yields exactly the rate function (2.3) for the family  $\tilde{V}_1^\varepsilon$ , in line with what one expects from Theorem 2.1.

### 3. PROOF OF THEOREM 2.1

The standard method to prove large deviations [12] is to first prove an upper bound for the lim sup, and then prove a lower bound for the lim inf, for the logarithmic probability of all Borel subsets of the real line. We prove here directly that the limit holds for all open intervals of the form  $(x, \infty)$  for  $x \in \mathbb{R}$ , which is clearly sufficient. For any  $t \geq 0$  and  $M \in \{X, V\}$ , define the rescaled cumulant generating function (cgf)  $\Lambda_M(\cdot, t)$  of the random variable  $M_t$  and its effective domain  $\mathcal{D}_t^M$  by

$$(3.1) \quad \Lambda_M(u, t) := t \log \mathbb{E} \left( e^{uM_t/t} \right), \quad \text{for all } u \in \mathcal{D}_t^M := \{u \in \mathbb{R} : |\Lambda_M(u, t)| < \infty\},$$

Define further  $\mathcal{D}_M := \cap_{t>0} \mathcal{D}_t^M$ . From [22, Part I, Section 6.3.4], we know that

$$(3.2) \quad \begin{aligned} \Lambda_M(u, t) &= -\frac{\mu t}{2} [g_t^M(u) + 2 \log f_t^M(u)], & f_t^X(u) &\equiv \cosh\left(\frac{d(\frac{u}{t})t}{2}\right) - \frac{g_t^X(u)}{td(\frac{u}{t})} \sinh\left(\frac{d(\frac{u}{t})t}{2}\right), \\ f_t^V(u) &\equiv 1 + \frac{u\xi^2}{2bt} (1 - e^{bt}), & g_t^X(u) &\equiv bt + \rho\xi u, & g_t^V(u) &\equiv 0, \end{aligned}$$

where  $d(u) \equiv [(b + \rho\xi u)^2 + u(1 - u)\xi^2]^{1/2}$ , so that the functions  $\Lambda_X(\cdot, t)$  and  $\Lambda_V(\cdot, t)$  are explicitly well defined on  $\mathcal{D}_t^X$  and  $\mathcal{D}_t^V$ . The pointwise limit functions  $\Lambda_M(u) := \lim_{t \downarrow 0} \Lambda_M(u, t)$ , for  $M \in \{X, V\}$ , read as follows:

**Lemma 3.1.** *The function  $\Lambda_M$  is null on  $\mathcal{D}_M$  and infinite outside, with  $\mathcal{D}_X = (u_-, u_+)$  and  $\mathcal{D}_V = (-\infty, 2/\xi^2)$ .*

*Proof.* The lemma follows from a simple yet careful analysis of the functions  $\Lambda_X(\cdot, t)$  and  $\Lambda_V(\cdot, t)$  together with their effective domains. Clearly here  $\mathcal{D}_t^V = (-\infty, u^V(t))$ , where  $u^V(t) \equiv 2bt/[\xi^2(e^{bt} - 1)]$  converges from above to  $2/\xi^2$  as  $t$  tends to zero. In [18], the authors showed that  $u_+(t)$  (resp.  $u_-(t)$ ) converges from above (resp. from below) to  $u_+$  (resp.  $u_-$ ) as  $t$  tends to zero, so that the limiting domain  $\cap_{t>0} \mathcal{D}_t^X$  is equal to  $(u_-, u_+)$ . The pointwise limits are then straightforward to prove.  $\square$

Define now the following functions on  $\mathcal{D}_M^o$ :

$$(3.3) \quad \begin{cases} f_0^X(u) &:= \cos(\bar{\rho}\xi u/2) - \frac{\rho}{\bar{\rho}} \sin(\bar{\rho}\xi u/2), & f_0^V(u) &:= 1 - \frac{u\xi^2}{2}, & f_1^V(u) &:= -\frac{bu\xi^2}{4}, \\ f_1^X(u) &:= \begin{cases} \frac{\rho(\xi + 2b\rho)}{4\bar{\rho}^2} \cos(\bar{\rho}\xi u/2) + \left(\frac{\xi + 2b\rho}{4\bar{\rho}} - \frac{\xi\rho + 2b}{2u\xi\bar{\rho}^3}\right) \sin(\bar{\rho}\xi u/2), & \text{if } u \neq 0, \\ -b/2, & \text{if } u = 0, \end{cases} \\ g_0^X(u) &:= \rho\xi u, & g_0^V(u) &\equiv 0. \end{cases}$$

**Lemma 3.2.** *For  $M \in \{X, V\}$ , the expansions  $f_t^M(u) = f_0^M(u) + f_1^M(u)t + \mathcal{O}(t^2)$  and  $g_t^M(u) = g_0^M(u) + \mathcal{O}(t)$  hold for all  $u \in \mathcal{D}_M^o$  as  $t$  tends to zero. They further hold uniformly on compacts.*

*Proof.* Let  $M = X$ , and define the quantities  $d_0 := \bar{\rho}\xi \operatorname{sgn}(u)$ ,  $d_1 := \frac{i(2\kappa\rho - \xi)\operatorname{sgn}(u)}{2\bar{\rho}}$ , where  $\operatorname{sgn}(u) = 1$  if  $u \geq 0$ , and  $-1$  otherwise; then for any  $u \in \mathcal{D}_X^o \setminus \{0\}$ ,  $d(u/t) = iud_0/t + d_1 + \mathcal{O}(t)$ , as  $t$  tends to zero, and hence

$$(3.4) \quad \begin{aligned} \frac{g_t^X(u)}{td(u/t)} &= \frac{\rho\xi u}{iud_0} + \frac{d_1\xi\rho - ibd_0}{d_0^2 u} t + \mathcal{O}(t^2), \\ \cosh\left(\frac{d(u/t)t}{2}\right) &= \cos\left(\frac{d_0 u}{2}\right) + \frac{id_1}{2} \sin\left(\frac{d_0 u}{2}\right) t + \mathcal{O}(t^2), \\ \sinh\left(\frac{d(u/t)t}{2}\right) &= i \sin\left(\frac{d_0 u}{2}\right) + \frac{d_1}{2} \cos\left(\frac{d_0 u}{2}\right) t + \mathcal{O}(t^2). \end{aligned}$$

The expansion for  $f_t^X$  in (3.2) for  $u \in \mathcal{D}_X^o \setminus \{0\}$  follows after using the asymptotics in (3.4) and some simplification. When  $u = 0$ , straightforward computations reveal that  $f_t^X(u) = 1 - bt/2 + \mathcal{O}(t^2)$ , in agreement with (3.2). Note that  $f_1^X$  is continuous at the origin. The expansions for  $f_t^V$  and  $g_t^M$  follow analogous arguments and the lemma follows. Uniform convergence of the sequences  $(g_t^M)_t$  and  $(f_t^V)_t$  is trivial. Uniform convergence on compacts of the sequence  $(f_t^X)_t$  holds as soon as  $\sup_{u \in \mathcal{D}_X^o} |f_t^M(u) - f_0^M(u) - f_1^M(u)t|$  converges to zero when  $t$  tends to zero, which is tedious but straightforward to prove.  $\square$

Consider now the (time-dependent) saddlepoint equation:

$$(3.5) \quad \partial_u \Lambda_M(u, t) = x, \quad x \in \mathcal{K}_M, t > 0.$$

The following lemma proves existence and uniqueness of the solution to this equation, as well as a small-time expansion. Let us first define the following functions on  $\mathcal{K}_M$ :

$$(3.6) \quad \alpha_0^X(x) := u_- \mathbf{1}_{\{x < 0\}} + u_+ \mathbf{1}_{\{x > 0\}}, \quad \alpha_0^V(x) := \frac{2}{\xi^2} \mathbf{1}_{\{x \geq 0\}},$$

**Lemma 3.3.** *For any  $x \in \mathcal{K}_M$ ,  $t > 0$ , Equation (3.5) admits a unique solution  $u_M^*(x, t) \in \mathcal{D}_t^M$ ; moreover, the expansion  $u_M^*(x, t) = \alpha_0^M(x) + \mathcal{O}(t)$  holds. For  $x = 0$ , Equation (3.5) also admits a unique solution  $u_M^*(0, t)$ , which converges to zero as  $t$  approaches zero.*

*Proof.* We first prove existence and uniqueness of the solution of the saddlepoint equation (3.5). Consider first the case  $M = V$ . Clearly, for any  $t > 0$ , the map  $\partial_u \Lambda_V(\cdot, t) : \mathcal{D}_t^V \rightarrow \mathbb{R}$  is strictly increasing and the image of  $\mathcal{D}_t^V$  by  $\partial_u \Lambda_V(\cdot, t)$  is  $\mathbb{R}_+$ . Thus, for any  $x > 0$ , (3.5) admits a unique solution  $u_V^*(x, t) = \frac{2t}{\xi^2} \left( \frac{b}{e^{bt}-1} - \frac{a}{x} \right)$ , which converges to  $\alpha_0^V(x)$ . Consider now the case  $M = X$ . We first start with the following claims, which can be proved using the convexity of the moment generating function and tedious computations.

- (i) For any  $t > 0$ , the function  $\partial_u \Lambda_X(\cdot, t) : \mathcal{D}_t^X \rightarrow \mathbb{R}$  is strictly increasing and maps  $\mathcal{D}_t^X$  to  $\mathbb{R}$ ;
- (ii) For any  $t > 0$ ,  $u_X^*(0, t) > 0$  and  $\lim_{t \downarrow 0} u_X^*(0, t) = 0$ , i.e. the unique minimum of  $\Lambda_X(\cdot, t)$  converges to zero;
- (iii) For each  $u \in \mathcal{D}_X^o$ ,  $\partial_u \Lambda_X(u, t)$  converges to zero as  $t$  tends to zero.

Now, choose  $x > 0$  (analogous arguments hold for  $x < 0$ ). It is clear from (i) that (3.5) admits a unique solution. Note further that (i) and (ii) imply  $u_X^*(x, t) > 0$ . Next we introduce the following condition.

**Condition A :** There exists  $t_1 > 0$  such that  $u_X^*(x, t) \in \mathcal{D}_X^o$  for all  $t < t_1$ .

Suppose condition A is not true and further assume that the sequence  $(u_X^*(x, t))_{t>0}$  does not converge to  $u_+^X$  as  $t$  tends to zero. Then there exists  $t_1^* > 0$  and  $\varepsilon > 0$  such that for all  $t < t_1^*$  we have  $u_X^*(x, t) \notin B(u_+^X, \varepsilon) := \{y \in \mathbb{R} : |y - u_+^X| < \varepsilon\}$ . But since  $\lim_{t \downarrow 0} \mathcal{D}_t^X = \mathcal{D}_X$ , this implies that our sequence must then satisfy condition A, which is a contradiction. Therefore  $u_X^*(x, t)$  converges to  $u_+^X$ . Next suppose that condition A is true. Again note that (i) and (ii) imply  $u_X^*(x, t) > 0$ . From (iii) there exists  $t_2 > 0$  such that the sequence  $(u_X^*(x, t))_{t>0}$  is strictly increasing as  $t$  goes to zero for  $t < t_2$ . Now let  $t^* = \min(t_1, t_2)$  and consider  $t < t^*$ . Then  $u_X^*(x, t)$  is bounded above by  $u_+$  (because  $u_X^*(x, t) \in \mathcal{D}_X^o$ ) and therefore converges to a limit  $L \in [0, u_+]$ . Suppose that  $L \neq u_+$ . Since  $s \mapsto u_X^*(x, s)$  is strictly increasing as  $s$  tends to zero (and  $s < t^*$ ), and  $\partial_u \Lambda_X(\cdot, t)$  is strictly increasing we have  $\partial_u \Lambda_X(u_X^*(x, t), t) \leq \partial_u \Lambda_X(L, t)$ ; Combining this and (iii) yields  $\lim_{t \downarrow 0} \partial_u \Lambda_X(u_X^*(x, t), t) \leq \lim_{t \downarrow 0} \partial_u \Lambda_X(L, t) = 0 \neq x$ , which contradicts the assumption  $x > 0$ . Therefore  $L = u_+$  and the first part of the lemma follows.

Given existence and uniqueness of the solution to the saddlepoint equation, we now prove the expansion stated in the lemma. In light of (3.2), the saddlepoint equation (3.5) can be written explicitly as

$$-\frac{\mu t}{2} [\partial_u g_t^M(u_M^*(x, t)) f_t^M(u_M^*(x, t)) + 2 \partial_u f_t^M(u_M^*(x, t))] = f_t^M(u_M^*(x, t)).$$

Using Lemma 3.2 in this equation and solving at each order yields the desired expansion.  $\square$

For  $M \in \{X, V\}$  and  $t > 0$ , introduce now a time-dependent change of measure by

$$(3.7) \quad \frac{d\mathbb{Q}_{x,t}^M}{d\mathbb{P}} := \exp \left( \frac{u_M^*(x, t) M_t}{t} - \frac{\Lambda_M(u_M^*(x, t), t)}{t} \right).$$

By Lemma 3.3,  $u_M^*(x, t)$  belongs to the interior of  $\mathcal{D}_t^M$ , and so  $|\Lambda_M(u_M^*(x, t))|$  is finite. Also  $d\mathbb{Q}_{x,t}^M/d\mathbb{P}$  is almost surely strictly positive and  $\mathbb{E}[d\mathbb{Q}_{x,t}^M/d\mathbb{P}] = 1$ . Therefore (3.7) is a valid measure change for all  $t > 0$  and  $x \in \mathcal{K}_M$ .



Define now the random variable  $Z_{x,t}^M := (M_t - x)$ , and denote its characteristic function in the  $\mathbb{Q}_{x,t}^M$ -measure (3.7) by  $\Phi_{x,t}^M(u) := \mathbb{E}^{\mathbb{Q}_{x,t}^M}(e^{iuZ_{x,t}^M})$ . We apply Fourier inversion methods to derive the tail probabilities of  $Z_{x,t}^M$  under the measure (3.7). It follows from Lemmas B.1 and B.2 that, for any  $x \in \mathcal{K}_M$ , the following estimate holds as  $t$  tends to zero:

$$(3.8) \quad \mathbb{E}^{\mathbb{Q}_{x,t}^M} \left[ \exp \left( -\frac{u_M^*(x,t)Z_{x,t}^M}{t} \right) \mathbf{1}_{\{Z_{x,t}^M \geq 0\}} \right] \mathbf{1}_{\{x > 0\}} = \mathcal{O}(t).$$

Finally, we have the following additional estimate.

**Lemma 3.4.** *For any  $x \in \mathcal{K}_M$ , there exists a constant  $C > 0$  such that the expansion*

$$\exp \left[ -\frac{xu_M^*(x,t)}{t} + \frac{\Lambda_M(u_M^*(x,t),t)}{t} \right] = C \exp \left( -\frac{\Lambda_M^*(x)}{t} \right) \mathcal{O}(t^{-\mu})$$

*holds as  $t$  tends to zero, with  $\Lambda_M^*$  in (2.2),(2.3).*

**Remark 3.5.** The constant  $C$  above can be computed explicitly as

$$C = \left( -\frac{\mu \partial_u f_0^M(\alpha_0^M(x))}{x} \right)^{-\mu} \exp \left( -x\alpha_1^M(x) - \frac{\mu}{2} g_0^M(\alpha_0^M(x)) \right).$$

*Proof.* From Lemma 3.3 and the characterisation of  $\Lambda^*$  in (2.2),(2.3), there exists  $C > 0$ , such that for small  $t$ ,

$$(3.9) \quad \exp \left( -\frac{xu_M^*(x,t)}{t} \right) = C \exp \left( -\frac{x\alpha_0^M(x)}{t} \right) (1 + \mathcal{O}(t)) = C \exp \left( -\frac{\Lambda_M^*(x)}{t} \right) (1 + \mathcal{O}(t)).$$

The definition of  $\Lambda_M$  in (3.2) and Lemma 3.3 yield  $e^{\Lambda_M(u_M^*(x,t),t)/t} = \mathcal{O}(t^{-\mu})$ , and the lemma follows.  $\square$

We now put all the pieces together. Using the time-dependent change of measure (3.7), we have for  $x > 0$ ,

$$\begin{aligned} \mathbb{P}(M_t \geq x) &= \mathbb{E} [\mathbf{1}_{\{M_t \geq x\}}] = \exp \left( \frac{\Lambda_M(u_M^*(x,t))}{t} \right) \mathbb{E}^{\mathbb{Q}_{x,t}^M} \left[ \exp \left( -\frac{u_M^*(x,t)M_t}{t} \right) \mathbf{1}_{\{M_t \geq x\}} \right] \\ &= \exp \left( -\frac{xu_M^*(x,t) - \Lambda_M(u_M^*(x,t),t)}{t} \right) \mathbb{E}^{\mathbb{Q}_{x,t}^M} \left[ \exp \left( -\frac{u_M^*(x,t)Z_{x,t}^M}{t} \right) \mathbf{1}_{\{Z_{x,t}^M \geq 0\}} \right], \end{aligned}$$

with  $Z_{x,t}^M$  defined on page 8. An analogous argument holds for probabilities  $\mathbb{P}(M_t \leq x)$  when  $x < 0$ , and Theorem 2.1 the follows from Lemma 3.4 and Equation(3.8).

## APPENDIX A. THE GÄRTNER ELLIS THEOREM

We provide here a brief review of large deviations and the Gärtner-Ellis theorem. For a detailed account of these, the interested reader should consult [12]. Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of random variables in  $\mathbb{R}$ , with law  $\mu_n$  and cumulant generating function  $\Lambda_n(u) \equiv \log \mathbb{E}(e^{uX_n})$ .

**Definition A.1.** The sequence  $X_n$  is said to satisfy a large deviations principle with speed  $n$  and rate function  $I$  if for each Borel measurable set  $E \subset \mathbb{R}$ ,

$$-\inf_{x \in E^\circ} I(x) \leq \liminf_{n \uparrow \infty} \frac{1}{n} \log \mathbb{P}(X_n \in E) \leq \limsup_{n \uparrow \infty} \frac{1}{n} \log \mathbb{P}(X_n \in E) \leq -\inf_{x \in \bar{E}} I(x).$$

Before stating the main theorem, we need one more concept:

**Definition A.2.** Let  $\Lambda : \mathbb{R} \rightarrow (-\infty, +\infty]$  be a convex function, and  $\mathcal{D}_\Lambda := \{u \in \mathbb{R} : \Lambda(u) < \infty\}$  its effective domain. It is said to be essentially smooth if

- The interior  $\mathcal{D}_\Lambda^\circ$  is non-empty;



- $\Lambda$  is differentiable throughout  $\mathcal{D}_\Lambda^\circ$ ;
- $\Lambda$  is steep:  $\lim_{n \uparrow \infty} |\Lambda'(u_n)| = \infty$  whenever  $(u_n)$  is a sequence in  $\mathcal{D}_\Lambda^\circ$  converging to a boundary point of  $\mathcal{D}_\Lambda^\circ$ .

Assume now that the limiting cumulant generating function  $\Lambda(u) := \lim_{n \uparrow \infty} n^{-1} \Lambda_n(nu)$ , exists as an extended real number for all  $u \in \mathbb{R}$ , and let  $\mathcal{D}_\Lambda$  denote its effective domain. Let  $\Lambda^* : \mathbb{R} \rightarrow \mathbb{R}_+$  denote its (dual) Fenchel-Legendre transform, via the variational formula  $\Lambda^*(x) \equiv \sup_{u \in \mathcal{D}_\Lambda} \{ux - \Lambda(u)\}$ . Then the following holds:

**Theorem A.3** (Gärtner-Ellis theorem). *If the origin lies in the interior of  $\mathcal{D}_\Lambda$  and if  $\Lambda$  is lower semicontinuous and essentially smooth, then the sequence  $(X_n)_n$  satisfies a large deviations principle with rate function  $\Lambda^*$ .*

## APPENDIX B. INVERSE FOURIER TRANSFORM REPRESENTATION

Let  $g(z) := \exp(-u_M^*(x, t)z/t) \mathbf{1}_{\{z \geq 0\}}$ . The main result of this appendix is the following representation:

**Lemma B.1.** *For every  $x > 0$ , there exists  $t_1^* > 0$  such that for all  $t < t_1^*$ :*

$$(B.1) \quad \mathbb{E}^{\mathbb{Q}_{x,t}^M} [g(Z_{x,t}^M)] = \frac{1}{2\pi} \int_{\mathbb{R}} \Phi_{x,t}^M(u) \left( \frac{u_M^*(x, t)}{t} + iu \right)^{-1} du.$$

The proof of Lemma B.1 proceeds in two steps: We first prove that the integrand in the right-hand side of Equality (B.1) belongs to  $L^1(\mathbb{R})$  (and hence the integral is well-defined), and then prove that this very equality holds. The first step is contained in the following lemma.

**Lemma B.2.** *We have  $\int_{\mathbb{R}} \left| \frac{\Phi_{x,t}^M(u)}{u_M^*(x, t)/t + iu} \right| du = \mathcal{O}(t)$  as  $t$  tends to zero. In particular, there exists  $t_0^* > 0$  such that  $\int_{\mathbb{R}} \left| \frac{\Phi_{x,t}^M(u)}{u_M^*(x, t)/t + iu} \right| du$  is finite for all  $t < t_0^*$ .*

*Proof.* From the change of measure (3.7) and the re-scaled cgf given in (3.1) we can compute

$$\log \Phi_{x,t}^M(u) = \log \mathbb{E}^{\mathbb{P}} \left( \frac{d\mathbb{Q}_{x,t}^M}{d\mathbb{P}} \exp(iuZ_{x,t}^M) \right) = -iux + \frac{1}{t} [\Lambda_M(iut + u_M^*(x, t), t) - \Lambda_M(u_M^*(x, t), t)].$$

Using the definition of  $\Lambda_M$  in (3.2) then yields

$$(B.2) \quad \Phi_{x,t}^M(u) = \left( \frac{f_t^M(iut + u_M^*(x, t))}{f_t^M(u_M^*(x, t))} \right)^{-\mu} \exp \left( -iux - \frac{2}{\mu} [g_t^M(u_M^*(x, t) + iut) - g_t^M(u_M^*(x, t))] \right).$$

Let now  $M = V$ . Using (B.2) we see that  $\Phi_{x,t}^V(u) = e^{-iux} (1 + iud_t)^{-\mu}$  where  $d_t := \frac{\xi^2 t (1 - e^{bt})}{2bt + u_V^*(x, t) \xi^2 (1 - e^{bt})}$ . It is easy to see that  $\lim_{t \downarrow 0} d_t = b\xi^2$ . The modulus is then given by  $|\Phi_{x,t}^V(u)| = (1 + u^2 d_t^2)^{-\mu/2}$  and hence for small enough  $t$  we have that  $|\Phi_{x,t}^V(u)| \leq D|u|^{-\mu}$  for some  $D > 0$ . Furthermore, we easily see that  $|(u_V^*(x, t) + iut)^{-1}| \leq 1/|u_V^*(x, t)|$  for all  $t$ , and hence we compute

$$\begin{aligned} \int_{\mathbb{R}} \left| \frac{\Phi_{x,t}^V(u)}{u_V^*(x, t)/t + iu} \right| du &= t \int_{\mathbb{R}} \left| \frac{\Phi_{x,t}^V(u)}{u_V^*(x, t) + iut} \right| du = t \int_{|u| \leq 1} \left| \frac{\Phi_{x,t}^V(u)}{u_V^*(x, t) + iut} \right| du + t \int_{|u| > 1} \left| \frac{\Phi_{x,t}^V(u)}{u_V^*(x, t) + iut} \right| du \\ &\leq t \frac{1}{u_V^*(x, t)} \left( 1 + D \int_{|u| > 1} \frac{du}{|u|^\mu} \right). \end{aligned}$$

The factor multiplying  $t$  in the last inequality is bounded for sufficiently small  $t$ , since  $u_V^*(x, t)$  converges to  $2/\xi^2$  as  $t$  tends to zero, and  $\mu > 1$ . The case  $M = X$  follows from analogous, yet tedious, computations.  $\square$

We now move on to the proof of Lemma B.1. We only look at the case  $M = V$ , the other case being completely analogous. We denote the convolution of two functions  $f, h \in L^1(\mathbb{R})$  by  $(f * h)(x) := \int_{\mathbb{R}} f(x - y)h(y)dy$ , and recall that  $(f * h) \in L^1(\mathbb{R})$ . For such functions, we denote the Fourier transform by  $(\mathcal{F}f)(u) := \int_{\mathbb{R}} e^{iux} f(x)dx$  and the inverse Fourier transform by  $(\mathcal{F}^{-1}h)(x) := \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} h(u)du$ . We have that

$$(B.3) \quad \mathcal{F}(g(z))(u) := \int_{\mathbb{R}} \exp\left(-\frac{u_V^*(x, t)z}{t} + iuz\right) \mathbf{1}_{\{z \geq 0\}} dz = \left(\frac{u_V^*(x, t)}{t} - iu\right)^{-1},$$

if  $u_V^*(x, t) > 0$ , which holds for small  $t$  since by Lemma 3.3  $u_V^*(x, t)$  converges to  $\alpha_0^V$  and  $\alpha_0^V > 0$ . We write

$$\mathbb{E}^{\mathbb{Q}_{x,t}^V} [g(Z_{x,t}^V)] = \int_{\mathbb{R}} q(x - y)p(y)dy = (q * p)(x),$$

with  $q(z) \equiv g(-z)$  and  $p$  denoting the density of  $V_t$ . On the strips of regularity ( $x > 0$ ) we know there exists  $t_0 > 0$  such that  $q \in L^1(\mathbb{R})$  for  $t < t_0$ . Since  $p$  is a density,  $p \in L^1(\mathbb{R})$ , and therefore

$$(B.4) \quad \mathcal{F}(q * p)(u) = \mathcal{F}q(u)\mathcal{F}p(u).$$

We note that  $\mathcal{F}q(u) \equiv \mathcal{F}g(-u) \equiv \overline{\mathcal{F}g(u)}$  and hence using (B.3)

$$(B.5) \quad \mathcal{F}q(u)\mathcal{F}p(u) \equiv e^{iux} \Phi_{x,t}^V(u) \left(\frac{u_V^*(x, t)}{t} + iu\right)^{-1},$$

since the complex conjugate of  $w^{-1}$  is equal to  $(\Re(w) - i\Im(w))^{-1}$ , for  $w \in \mathbb{C}$ . Thus by Lemma B.2 there exists an  $t_1 > 0$  such that  $\mathcal{F}q\mathcal{F}p \in L^1(\mathbb{R})$  for  $t < t_1$ . By the inversion theorem [30, Theorem 9.11] this then implies from (B.4) and (B.5) that for  $t < \min(t_0, t_1)$ :

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_{x,t}^V} [g(Z_{x,t}^V)] &= (q * p)(x) = \mathcal{F}^{-1}(\mathcal{F}q(u)\mathcal{F}p(u))(x) \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \mathcal{F}q(u)\mathcal{F}p(u)du = \frac{1}{2\pi} \int_{\mathbb{R}} \Phi_{x,t}^V(u) \left(\frac{u_V^*(x, t)}{t} + iu\right)^{-1} du. \end{aligned}$$

## REFERENCES

- [1] M. Abramowitz and I. Stegun. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York: Dover Publications, 1972.
- [2] P. Baldi and L. Caramellino. General Freidlin-Wentzell large deviations and positive diffusions. *Statistics & Probability Letters*, 81: 1218-1229, 2011.
- [3] B. Bercu, L. Coutin and N. Savy. Sharp large deviations for the fractional Ornstein-Uhlenbeck process. *SIAM Theory of Probability and its Applications*, 55: 575-610, 2011.
- [4] B. Bercu, F. Gamboa and M. Lavielle. Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM PS*, 4: 1-24, 2000.
- [5] B. Bercu and A. Rouault, Sharp large deviations for the Ornstein-Uhlenbeck process, *SIAM Theory of Probability and its Applications*, 46: 1-19, 2002.
- [6] W. Bryc. Large deviations by the asymptotic value method. *Diffusion processes and related problems in analysis*, 1. Boston: Birkhäuser, 1990.
- [7] W. Bryc and A. Dembo. Large deviations for quadratic functionals of Gaussian processes. *Journal of Theoretical Probability*, 10: 307-332, 1997.
- [8] H. Comman. Differentiability-free conditions on the free-energy rate function implying large deviations *Conuantes Mathematici*, 1(2): 181-196, 2009.
- [9] G. Conforti, S. De Marco and J-D. Deuschel. On small-noise equations with degenerate limiting system arising from volatility models. Large Deviations and Asymptotic Methods in Finance (Editors: P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, J. Teichmann). Springer Proceedings in Mathematics and Statistics, 110, 2015.

- [10] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* 736: 5-23. Colloque consacré à la théorie des probabilités 3. Hermann, Paris, 1938.
- [11] A. Dembo and O. Zeitouni. Large deviations via parameter dependent change of measure and an application to the lower tail of Gaussian processes. In E. Bolthausen, M. Dozzi and F. Russo, editors, *Progress in Probability*, 36: 111-121. Birkhäuser, Basel, Switzerland, 1995.
- [12] A. Dembo and O. Zeitouni. Large deviations techniques and applications. Jones and Bartlet Publishers, Boston, 2nd Edition, 1998.
- [13] J-D. Deuschel, P.K. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility, Part II: Applications. *Communications on Pure and Applied Mathematics*, 67(2): 321-350, 2014.
- [14] C. Donati-Martin, A. Rouault, M. Yor and M. Zani. Large deviations for squares of Bessel and Ornstein-Uhlenbeck processes. *Probability Theory and Related Fields*, 129: 261-289, 2004.
- [15] K.R. Duffy and A. Sapozhnikov. The LDP for the on/off Weibull sojourn process. *Journal of Applied Probability*, 45(1): 107-117, 2008.
- [16] P. Dupuis and R.S. Ellis. Large deviations for Markov processes with discontinuous statistics, II: random walks. *Probability Theory and Related Fields*, 91: 153-194, 1992.
- [17] P. Eichelsbacher. Some Aspects in Large Deviations. Preprint, [citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.55.6507](http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.55.6507).
- [18] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the Heston model. *International Journal of Theoretical and Applied Finance*, 12(6), 861-876, 2009.
- [19] J. Gatheral. The Volatility Surface: A Practitioner's Guide. John Wiley & Sons, 2006.
- [20] A. Jacquier and P. Roome. The small-maturity Heston forward smile. *SIAM Journal on Financial Math.*, 4(1): 831-856, 2013.
- [21] A. Jacquier and P. Roome. Asymptotics of forward implied volatility *SIAM Journal on Financial Math.*, 6(1): 307-351, 2015.
- [22] M. Jeanblanc, M. Yor and M. Chesney. Mathematical Methods for Financial Markets. Springer, 2009.
- [23] Y.W. Hiang and L.M. Wu. Large Deviations for Empirical Measures of Not Necessarily Irreducible Countable Markov Chains with Arbitrary Initial Measures. *Acta Mathematica Sinica*, 21(6): 1377-1390, 2005.
- [24] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus. Springer-Verlag, 1997.
- [25] S. Karlin and H. Taylor. A Second Course in Stochastic Processes. Academic Press, 1981.
- [26] G. Leoni. A First Course in Sobolev Spaces. Graduate Studies in Mathematics, 105. American Mathematical Society, Cambridge, 2009.
- [27] G.L. O'Brien and W. Vervaat. Compactness in the theory of large deviations. *Stochastic Processes and their Applications*, 57: 1-10, 1995.
- [28] F.W.J. Olver. Asymptotics and Special Functions. A.K. Peters, Wellesley, MA, 1997.
- [29] H. Pham. Some methods and applications of large deviations in finance and insurance. Paris-Princeton Lecture notes in mathematical Finance, Springer Verlag, 2007.
- [30] W. Rudin. Real and complex analysis, third edition. McGraw-Hill, 1987.
- [31] D. Williams. Probability with Martingales. Cambridge University Press, 1991.

CMAP, ECOLE POLYTECHNIQUE PARIS

*E-mail address:* `demarco@cmap.polytechnique.fr`

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON

*E-mail address:* `a.jacquier@imperial.ac.uk`

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON

*E-mail address:* `p.roome11@imperial.ac.uk`